# Conversion of a database to SDS Techniques

# Data Mining

- **Create a correlation matrix of the origin database schema TO the Spatial Data Standards (SDS)**
  - **Entity type**
    - **Table**
      - **Attribute**
        - **Domain**

# Correlation Matrix
## Entity Type

| Original | | SDS | | | |
|---|---|---|---|---|---|
| FEATURES | TABLE | CLASS | ENTITY TYPE NAME | DISCRIMINATOR | TABLE |
| sediment sampling site | bed_sediment | lfbth | sediment_sample_site | Bottom,Suspended | lfbthsmp |
| water quality sampling site | environ_samples | ehcha | surface_water_quality_monitoring_station_point | | ehchaswm |
| major inflow structure | pollution | utsto | storm_sewer_headwall_point | | utstohdw |
| major outflow structure | pollution | utsto | storm_sewer_headwall_point | | utstohdw |
| camp site | camp_site | imrec | camping_site | | imreccmp |
| vessel pollution incident | v_poll_incid | ehpol | pollution_source_point | | ehpolpsp |
| facility pollution incident | f_poll_incid | ehpol | point_source_pollution | | ehpolpsp |
| pot HTRW | pot_HTRW_site | ehsit | hazards_potential_concern_site | | ehsitaoc |
| pot HTRW site | pot_HTRW_site | ehsit | hazards_potential_concern_site | | ehsitaoc |
| superfund site | superfund_site | ehsit | superfund_site | | ehsitepa |
| suspended sediment site | bed_sediment | lfbth | suspended_sediment_sample_site | | lfbthsus |
| national wetland inv | wetland_inv | hywet | wetland_area | | hywetlnd |
| aquatic habitat | aquatic_habitat | ecology_habitat | aquatic_habitat_area | | echabaqv |
| benthos sampling site | eco_samples | flmgt | flora_sample_site | | flmgtsmp |
| city/county wild mgt area | wildl_mgt_area | famgt | government_wildlife_management_area | CITY/COUNTY | famgtgov |
| federal wild mgt area | wildl_mgt_area | famgt | government_wildlife_management_area | FEDERAL | famgtgov |
| fish sampling site | eco_samples | fapis | pisces_habitat_sample_point | | famgtsmp |
| land cover | land_cover | lscnd | land_cover_area | | lscndlcv |
| least tern site | least_tern | fahab | fauna_special_habitat_sample_point | ENDANGERED | fahabspc |
| plankton sampling site | eco_samples | ecology_habitat | plankton_sampling_site | | echabplk |
| species habitat | species_habitat | ecology_habitat | ecology_habitat_site | | echabars |
| species site | species site | ecology_habitat | ecology_species_site | | echabars |
| state wild mgt area | wildl_mgt_area | famgt | government_wildlife_management_area | STATE | famgtgov |
| vegetation sampling site | eco_samples | flmgt | flora_sample_site | | flmgtsmp |
| wildlife sampling site | eco_samples | fagen | fauna_general_habitat_sample_point | | famgtsmp |

# Correlation Matrix
## Table/Attribute

| Feature | Table | Attribute | SDS Feature (Entity Type N | SDS Table | SDS Attribute |
|---|---|---|---|---|---|
| | | value_of_contents | | | value_cont |
| | | sname (*) | | | structname |
| | | owner_id | | | owner_id |
| elevators | Structures | mslink | grain_elevator_site | immacgel | datalink |
| | | mapid | | | map_id |
| | | river_id | | | |
| | | river_mile | | | sur_crs_id |
| | | sarea | | | area_size |
| | | slab_elevation | | | |
| | | no_of_floors | | | |
| | | material_type | | | str_mat_d |
| | | usage | | | str_use_d |
| | | stype | | | str_type_d |
| | | value_of_struct | | | a_cost |
| | | value_of_contents | | | value_cont |
| | | sname (*) | | | structname |
| | | owner_id | | | owner_id |
| factory | Structures | mslink | structure_existing_site | bggenstr | datalink |
| | | mapid | | | map_id |
| | | river_id | | | sur_crs_id |
| | | river_mile | | | river_mile |
| | | sarea | | | area_size |
| | | slab_elevation | | | floor_elev |
| | | no_of_floors | | | no_levels |
| | | material_type | | | str_mat_d |
| | | usage | | | str_use_d |
| | | stype | | | str_type_d |
| | | value_of_struct | | | a_cost |
| | | value_of_contents | | | value_cont |
| | | sname (*) | | | structname |
| | | owner_id | | | owner_id |

# Correlation Matrix

## Domain

| | | | | | | |
|---|---|---|---|---|---|---|
| const dredging-existing | Const_dredging | | lfbthsub | | | |
| bendway weir sys-existing | bweir_sys_mp | | imerobdr | | | |
| dike system existing | dike_sys_mp | | imfdcdkr | | | |
| foreshore-existing | Foreshore_mp | | imerofrr | | | |
| revetment-existing | revet_mp | rtype | imerorvr | rev_type_d | d_revtyp | A) ACM |
| | | | | | | F) Stone fill |
| | | | | | | P) Stone paving |
| | | | | | | T) Trenchfill |
| | | | | | | O) Other |
| | | rdirect | | rev_ori_d | d_revori | U) Upstream extension |
| | | | | | | D) Downstream extension |
| | | | | | | G) filling a gap between existing revetments |
| | | | | | | T) not connected to other revetments |
| | | | | | | O) other |
| disposal_area | disposal_area | | lsmgtdsp | | | |
| dredged_area | dredged_area | disposal_type | lfbthsub | met_disp_d | d_mthdsp | CHANNEL |
| | | | | | | BANK |
| | | | | | | DISPOSAL AREA |
| | | discharge_side | | dschsd_d | d_dschsd | RIGHT |
| | | | | | | LEFT |

# Components

- Document the original database and project remaining GIS elements:
  - Software (GIS application & Database)
  - Hardware
  - Procedures

# What happened to my original data?

- Every category and feature will be converted to a corresponding SDS entity set and entity type
  - this is a MAJOR change that needs to be totally comprehended !!!
- All CADD files will be modified using SDS naming conventions and validation rules
  - this impacts any customized tools !!!

# Conversion Concerns

- One to many and many to one splits in the database tables
  - ✳ where do all my attributes fit??
- CADD files linkages
  - ✳ ALL CADD file names and contents will change - how do I determine where each feature will go??

# Technique

- Build a database “shell” using “Original” Filter

- Create custom productivity tools to assist in conversion process:
  - ✳ CADD file commands to change linkages
  - ✳ checker Visual Basic tools to validate quality of database
  - ✳ ACCESS queries to translate database

- Separate the conversion process into 2 parts:
  - ✳ Attribution
  - ✳ Graphics

# Technique

ATTRIBUTION:

- Determine the tables and attributes which contain data - these are the **ONLY** ones that need to be converted
  - ✳ converting data types, domain tables, and building relationship links
- Verify all features and CADD file usage using a custom checker tool

# Technique

GRAPHICS:

- Determine the CADD files that need to be converted using a custom application
- Move all graphic features to the correct SDS CADD file
  - level and symbology (color, line style, weight) may change
- Update linkages in all CADD files
  - The user tables' mslink values will be maintained

# Documentation

✦Two reports should be generated detailing:

- ✳ all quality control checks done before and after conversion
- ✳ the entire conversion process

# Summary of Changes

- Most or all Categories & Features may change

- Most or all CADD filenames may change

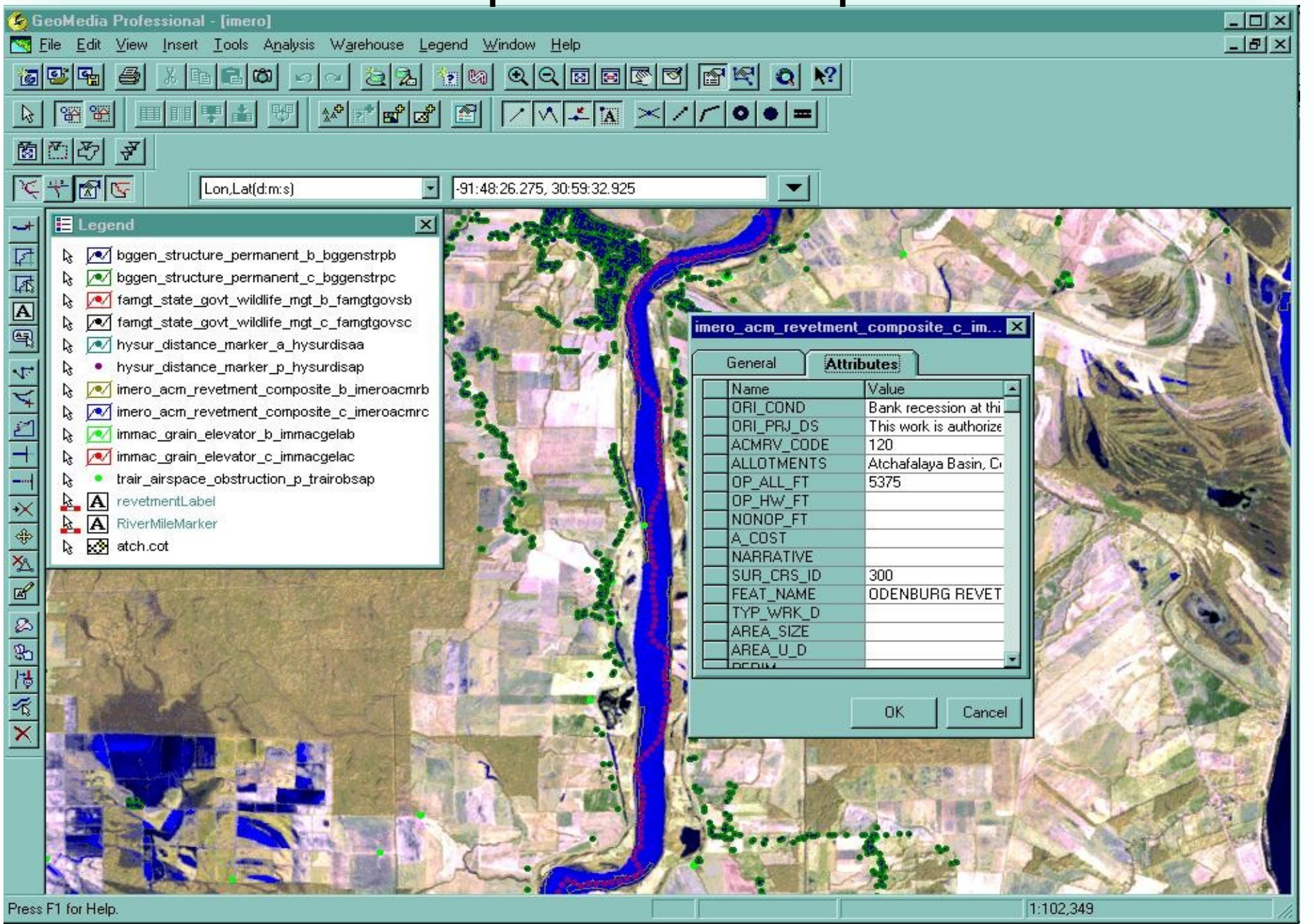
# Impact of Changes

- Application impacts:
  - reference files
  - Canned procedures
  - any other applications

# Lessons Learned

- Cost calculation:
  - ✳Correlation matrix – 60% of total cost
  - ✳Data conversion - costs approx $300 per feature (entity type) to convert tables and CADD files (40% of total cost)
- Correct, clean original data (database and graphics) enables the conversion to go more smoothly
- Thorough knowledge of SDS is a **MUST** in conversion process

# A SDS-compliant example dataset



# Questions ??


**Nancy Towne**

**601-634-3181 Voice**
**601-634-4584 Fax**
**E-mail : townen@wes.army.mil**